

PRIOR ART INFORMATION LIST

Your case No.	
Our case No.	2001FJ514

Inventor, Patent number, Country, Author, Title, Number of Document	Issue date	Concise Explanation of the Relevance (indication of page, column, line, figure of the relevant portion)
JP-A-11-219407	August 10, 1999	U.S. Application No. 09/193,868 (Your Docket: 826.1519)

**BEST AVAILABLE COPY** **BEST AVAILABLE COPY**

(19)日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11)特許出願公開番号

特開平11-219407

(43)公開日 平成11年(1999)8月10日

(51)Int.Cl.<sup>a</sup>  
G 0 6 K 9/20  
G 0 6 T 7/60  
G 0 6 K 9/32

識別記号  
3 4 0  
3 2 0

F I  
G 0 6 K 9/20  
G 0 6 F 15/70

3 4 0 K  
3 2 0 K  
9/32  
3 5 0 H

審査請求 未請求 請求項の数21 OL (全 21 頁)

(21)出願番号 特願平10-146926

(22)出願日 平成10年(1998)5月28日

(31)優先権主張番号 特願平9-327625

(32)優先日 平9(1997)11月28日

(33)優先権主張国 日本 (J P)

(71)出願人 000005223

富士通株式会社  
神奈川県川崎市中原区上小田中4丁目1番  
1号

(72)発明者 藤本 克仁

神奈川県川崎市中原区上小田中4丁目1番  
1号 富士通株式会社内

(72)発明者 鎌田 洋

神奈川県川崎市中原区上小田中4丁目1番  
1号 富士通株式会社内

(72)発明者 黒川 浩司

神奈川県川崎市中原区上小田中4丁目1番  
1号 富士通株式会社内

(74)代理人 弁理士 小笠原 吉義 (外2名)

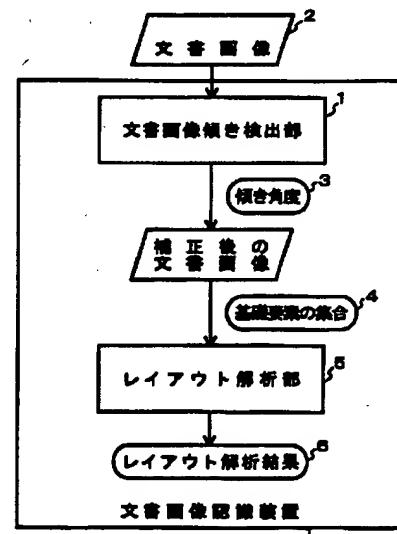
(54)【発明の名称】 文書画像認識装置および文書画像認識プログラムの記憶媒体

(57)【要約】

【課題】文書画像の傾き検出および行と段の抽出機能を持つ文書画像認識装置に関し、横行と縦行が混在する文書画像の傾き検出および行・段の抽出を高速かつ高精度に行うことができるようすることを目的とする。

【解決手段】文書画像傾き検出部1は、文書画像を参照して文字を構成する成分の集合を抽出し、これをもとに横および縦の行候補を抽出し、それらの各々の行候補に対して信頼度を推定し、推定した行信頼度に基づいて確度の高い行の集合を抽出する。その結果をもとに確度の高い行の集合の文字成分の配置を用いて傾きを推定する。レイアウト解析部5は、文書画像を構成する基礎要素の集合から行方向の確定した行集合を抽出し、行同士の関連付けによる段抽出と段を制約とした行抽出とを相互に実行して行と段を抽出する。

本発明の基本構成例



## 【特許請求の範囲】

【請求項1】 電子化された文書画像の傾きを検出し、その傾きを補正して文書画像を認識する文書画像認識装置において、傾き検出対象の文書画像を参照し、文字を構成する成分の集合を抽出する文字成分抽出手段と、前記抽出した文字成分の集合を参照し、横および縦の行候補を抽出する行候補抽出手段と、各々の行候補に対して信頼度を推定する行信頼度推定手段と、前記推定した行信頼度に基づいて確度の高い行の集合を抽出する行抽出手段と、確度の高い行の集合の文字成分の配置を用いて傾きを推定する傾き推定手段とを備えることを特徴とする文書画像認識装置。

【請求項2】 前記文字成分抽出手段は、文書画像中の黒画素の連結成分を文字成分の候補として、または文書画像中の黒画素の各連結成分に対する外接矩形を求める、これらの外接矩形の重なり矩形を文字成分の候補として用いることを特徴とする請求項1記載の文書画像認識装置。

【請求項3】 前記行信頼度推定手段は、各々の行候補に対して行らしさを定量的に計算し、それぞれの文字成分外接矩形に対して、それが属する横行候補および縦行候補の行らしさを比較し、より行らしい方向を文字方向として求め、それぞれの行候補に対してその行候補に属する文字成分のうちその行候補の行方向と一致する文字方向を持つものの比率をその行候補の信頼度とすることを特徴とする請求項1記載の文書画像認識装置。

【請求項4】 前記傾き推定手段は、確度の高い行に属する文字成分外接矩形の基準辺の座標値から、複数直線の当てはめによる最小自乗法を用いて傾きの推定を行うことを特徴とする請求項1記載の文書画像認識装置。

【請求項5】 認識対象となる文書画像の傾きを検出し、その傾きを補正して文書画像を認識する文書画像認識装置に用いるプログラムを格納したプログラム記憶媒体であって、入力した文書画像を参照し文字を構成する成分の集合を抽出する文字成分抽出処理と、前記抽出した文字成分の集合を参照し、横および縦の行候補を抽出する行候補抽出処理と、各々の行候補に対して信頼度を推定する行信頼度推定処理と、前記推定した行信頼度に基づいて確度の高い行の集合を抽出する行抽出処理と、確度の高い行の集合の文字成分の配置を用いて傾きを推定する傾き推定処理とを計算機に実行させるプログラムを格納したことを特徴とする文書画像認識プログラムの記憶媒体。

【請求項6】 電子化された文書画像の認識にあたって、文書画像のレイアウトを解析する文書画像認識装置において、文書を構成する基礎要素集合から行方向の確定した行集合を抽出する基本行抽出手段と、行同士の関連付けによる段抽出と、段を制約とした行抽出とを相互に実行して、行と段を抽出する行・段相互抽出手段とを備えることを特徴とする文書画像認識装置。

【請求項7】 前記基礎要素集合として、文書画像中の黒画素連結成分または文書画像中の黒画素の連結成分の外接矩形の重なり矩形を用いることを特徴とする請求項6記載の文書画像認識装置。

【請求項8】 前記基本行抽出手段は、前記基礎要素集合の関連付けをもとに、横方向および縦方向の行方向を持つ行要素を生成する行要素生成手段と、各々の行要素に対して信頼度を推定し、信頼度に基づいて行要素を淘汰する行要素淘汰手段と、整合性のとれた行および文字集合を得るための行要素整合手段とを備えることを特徴とする請求項6記載の文書画像認識装置。

【請求項9】 前記基礎要素集合の関連付けを、要素間の距離などの近接性、または要素の大きさもしくは要素間隔がほぼ均等であるなどの同質性に基づいて行うこととする請求項8記載の文書画像認識装置。

【請求項10】 前記行要素淘汰手段は、各々の行要素に対して行信頼度を定量的に計算し、それぞれの文字要素の属する横方向行要素および縦方向行要素の行信頼度の高い方向を文字要素の方向として、それぞれの行要素に対してその行候補に属する文字要素のうち、その行要素の行方向と一致する文字方向を持つものの比率をその行要素の信頼度とすることを特徴とする請求項8記載の文書画像認識装置。

【請求項11】 前記行要素の信頼度が、その行要素に含まれている文字要素の数が一定数以上あるという行の長さによる量、文字要素同士が密に配置されているという近接性による量、または文字要素の大きさもしくは間隔がほぼ同じくらいであるなどの同質性による量を用いて定量化されることを特徴とする請求項8記載の文書画像認識装置。

【請求項12】 前記行・段相互抽出手段は、互いに関連付けられた段同士の統合を行う段要素統合手段と、整合性のとれた段および行集合を得るための段要素整合手段と、整合性のとれた行および文字集合を得るための行要素整合手段とを備えることを特徴とする請求項6記載の文書画像認識装置。

【請求項13】 前記行・段相互抽出手段は、段同士の統合を繰り返し実行することを特徴とする請求項6記載の文書画像認識装置。

【請求項14】 前記行・段相互抽出手段における段同士の関連付けを、行方向、行垂直方向とそれぞれ変化させて繰り返し実行することを特徴とする請求項12記載の文書画像認識装置。

【請求項15】 前記行・段相互抽出手段における段同士の関連付けの条件を、関連付けの対象となる段要素の内部構造により動的に変化させることを特徴とする請求項6記載の文書画像認識装置。

【請求項16】 前記行・段相互抽出手段における段同士の関連付けを、要素間の距離などの近接性、または要素の大きさもしくは要素間隔がほぼ均等であるなどの同

質性に基づいて行うことを特徴とする請求項12記載の文書画像認識装置。

【請求項17】前記行・段相互抽出手段における段同士の関連付けに用いるパラメータを、処理回数に応じて変化させることを特徴とする請求項12記載の文書画像認識装置。

【請求項18】前記行・段相互抽出手段における関連付けの対象となる段要素の内部構造として、段の外接矩形の大きさ、段に含まれる行数もしくは文字数、または段の外接矩形の位置を用いることを特徴とする請求項12記載の文書画像認識装置。

【請求項19】前記基本行抽出手段における行要素整合手段または前記行・段相互抽出手段における行要素整合手段は、行要素に属する文字要素の集合から、その行の標準の文字の幅、高さ、間隔などの行の属性となる情報を計算して保持しておき、要素間の関連付けのために活用することを特徴とする請求項8または請求項12記載の文書画像認識装置。

【請求項20】前記行・段相互抽出手段における段要素整合手段は、段要素に属する行要素の集合から、その段の標準の行の幅、高さ、間隔などの段の属性となる情報を計算して保持しておき、要素間の関連付けのために活用することを特徴とする請求項12記載の文書画像認識装置。

【請求項21】電子化された文書画像の認識にあたって、文書画像のレイアウトを解析する文書画像認識装置に用いるプログラムを格納したプログラム記憶媒体であって、文書を構成する基礎要素集合から行方向の確定した行集合を抽出する基本行抽出処理と、行同士の関連付けによる段抽出と、段を制約とした行抽出とを相互に実行して、行と段を抽出する行・段相互抽出処理とを計算機に実行させるプログラムを格納したことを特徴とする文書画像認識プログラムの記憶媒体。

#### 【発明の詳細な説明】

##### 【0001】

【発明の属する技術分野】本発明は、例えばイメージキャナで読み取った文書やファクシミリ装置で受信した文書などの文書画像の傾きを検出してその傾きを補正し、また文書画像中文字行および段を抽出して文書画像を認識するための文書画像認識装置および文書画像認識プログラムの記憶媒体に関する。

【0002】文書光学式読み取り(OCR)エンジンの対象文書の拡大のためには、新聞等に見られるような縦行横行混在文書のレイアウト解析機能の搭載が必須である。本発明は、この縦行横行混在文書のレイアウト解析のために必要な要素技術として、文書画像の傾き補正のための傾き検出処理、さらに文書画像認識のための行抽出処理および段抽出処理の新たな技術を提供するものである。

##### 【0003】

#### 【従来の技術】(1)文書画像傾き検出

一般の印刷文書を読み取るためには、まず、イメージキャナなどの画像入力装置を用いて文書画像として取り込む必要があるが、通常は原稿設置の際に多少傾きが生じる。電子ファイリングあるいは文書認識の使用に供するためには、その文書画像の傾き量を検出し、補正する必要がある。

【0004】従来の傾き検出技術においては、文書画像中の主要な構成要素であるテキスト領域では文字が規則正しく並んでいることを利用している。例えば、中野らは「文書画像の傾き補正のための一方式」と題して電子情報通信学会論文誌D, vol.J69-D, No.1, pp.1833-1834に記載されているように、文字列の基準線がほぼ一定に存在することに着目し、文字ブロックの下端座標値をハフ変換し、ハフ空間上でのピーク値を検出することによって文字列の傾きを推定する第1の方式を提案している。

【0005】また、水野らは「文書画像傾き検出装置」と題して特開平7-192085号公報に記載されているように、文字を構成する連結成分を抽出し、近接する連結成分同士を結合し仮の文字行を生成し、仮の文字行に接する直線を求ることによって文字列の傾きを推定する第2の方式を提案している。

【0006】また、斎藤らは「文書傾き補正装置」と題して特開平2-170280号公報に記載されているように、傾き角度θを順次変更しながら文書画像を仮に補正し、補正画像中の全黒画素を含む外接矩形の面積が最小となる角度θを傾き角度として求める第3の方式を提案している。

##### 【0007】(2)レイアウト解析(行・段の抽出)

従来、OCR装置などにおいて、横行縦行の混在した文書画像中の文字行および段を抽出する方法としては以下のようない方法が提案されている。

【0008】例えば、辻本らは、「文書画像処理装置」と題して特開平1-183783号公報に記載されているように、入力文書の文字行をある方向に射影し、その周辺分布を作成することにより、入力文書から自動的にその段組を決定する第4の方式を提案している。

【0009】また、水谷らは、「文書画像処理装置」と題して特開平5-174179号公報に記載されているように、入力文書の構成要素の空白領域を用いて段抽出を行う第5の方式を提案している。

【0010】また、平本らは、「文字行抽出方法および装置」と題して特開平10-31716号公報に記載されているように、文字行方向が混在しており、文字の大きさやピッチの異なる領域が混在する文書からの文字行抽出を行う第6の方式を提案している。

##### 【0011】

【発明が解決しようとする課題】一般の印刷文書には縦書き、横書きが混在した文書が多く、文書認識する際に

は、文字行および段落を適切に抽出する必要がある。

【0012】しかしながら、上記の従来の方式では、以下のような問題がある。

(1) 文書画像傾き検出に関する問題

上記第1の方式では、行方向一定を前提としているため、新聞のように横行と縦行の混在する文書に適用することができない。また、行方向一定の文書に対しても、すべての文字の下端が基準線上に存在するとは限らないため、誤差が含まれることが避けられない。さらに、ハフ変換処理には膨大な計算量を要するという問題がある。

【0013】また、上記第2の方式では、新聞のように横行と縦行の混在する文書の場合に、縦行の段組から誤って横の仮行を抽出する場合が発生するために、傾き推定の大きな誤差を生じてしまう可能性がある。

【0014】また、上記第3の方式では、横行縦行混在文書の傾き検出を目的として設計されてはいるが、文書画像の黒画素を含む外接矩形の面積という少量の情報をもとに角度検出を行っているため、検出した傾きの精度が不安定であるという問題点がある。また、画像自体を回転させて矩形領域を抽出する処理を何度も繰り返す必要があるため、膨大な計算量となる。

【0015】(2) レイアウト解析に関する問題

文字行と段を抽出するための上記第4の方式では、文字行をあらかじめ抽出しておき、それをもとに段抽出を行うため、段中の文字行が細分化されている不定形の段が細分化されてしまうという問題がある。

【0016】また、第5の方式では、空白領域を用いて段を抽出するために、行間より段間が狭い箇所がある文書の場合には誤った段抽出がされてしまうおそれがある。この問題は、縦行横行の文章が密に混在するような文書画像では大きな問題となる。例えば図26に示す新聞紙面の左上側の写真領域の下に矩形の枠で示すように、縦書の記事本文と横書の写真のキャプションの間が狭い文書画像の場合には、これらを一つの段落として、横書キャプションの各行の文字を縦書記事本文の先頭2文字と認識してしまうという問題が起こる。

【0017】また、第6の方式では、高精度な行抽出を行う前処理として段領域の抽出を行っているため、段中の文字行が細分化されている不定形の段が細分化されてしまい、結果として誤った行抽出をしてしまうという問題がある。

【0018】すなわち、従来技術では、①(基礎要素集合)→行抽出処理→段抽出処理→(レイアウト解析結果)、または②(基礎要素集合)→段抽出処理→行抽出処理→(レイアウト解析結果)のいずれかの手順をとり、ボトムアップ処理あるいはトップダウン処理を基本としている。これらの従来技術は、どれも行抽出処理、段抽出処理を独立したものと捉え、逐次的に両者を実行することにより、行および段抽出を達成しようとしている。

るところが共通しており、ここに問題点発生の主要な原因が存在する。

【0019】このような技術的背景を踏まえて、本発明は、横行縦行が混在した複雑な文書構造を持つ文書画像を認識するような場合にも、横行縦行混在文書の傾き検出を高速かつ高精度に行うことができ、また行および段の抽出を高精度に行うことができる文書画像認識装置を提供することを目的とする。

【0020】

10 【課題を解決するための手段】図1は、本発明の基本構成例を示す。縦行横行が混在することがある認識対象の文書画像2が、文書画像認識装置100に入力されると、まず、文書画像傾き検出部1は、傾き補正のための傾き角度3を検出する。レイアウト解析部5は、この傾き角度3によって補正した補正後の文書画像から抽出した基礎要素の集合4について、基本行を抽出し、行および段抽出を相互に繰り返し行うことで、横行縦行が混在する文書画像2の段構造を認識し、レイアウト解析結果6を出力する。

20 【0021】(1) 文書画像傾き検出

文書画像傾き検出部1は、文書画像2を参照し文字を構成する成分の集合を抽出する文字成分抽出手段と、文字成分の集合を参照し横および縦の行候補を抽出する行候補抽出手段と、各々の行候補に対して信頼度を推定する行信頼度推定手段と、行信頼度に基づいて確度の高い行の集合を抽出する行抽出手段と、確度の高い行の集合の文字成分の配置を用いて傾きを推定する傾き推定手段とを備える。

【0022】文字成分抽出手段で抽出した文字成分の集合を用いて後段の処理を行い、文書画像の回転など画像を直接操作しないため、計算量を大幅に削減できる。行候補抽出手段によって横方向および縦方向の行の候補を抽出し、行信頼度推定手段および行抽出手段により確度の高い横および縦方向の行の集合を抽出するため、横行縦行混在文書に対しても傾き検出を行うことができる。さらに、傾き推定手段では、確度の高い行を構成する文字成分のみを用いて傾きを推定するために、ノイズに対して強く高精度の傾き推定を実現することができる。

【0023】本発明を用いると、従来技術では難しかった横行縦行混在文書の傾き検出を、高速かつ高精度に行うことができるようになる。

(2) レイアウト解析(行・段の抽出)

レイアウト解析部5は、文書を構成する基礎要素の集合4から方向の確定した行集合を抽出する基本行抽出手段と、行同士の関連付けによる段抽出および段を制約とした行抽出を相互に実行して行と段とを抽出する行・段相互抽出手段とを備える。ここで入力となる基礎要素の集合4としては、例えば文書画像中の黒画素連結成分、あるいは文書画像中の黒画素の連結成分の外接矩形の重なり矩形を用いることができる。

【0024】従来技術との大きな違いは、従来技術では行抽出処理後に段抽出処理、または段抽出処理後に行抽出処理というように、行抽出処理と段抽出処理とを独立に行っていったのに対し、本発明では、行の抽出と段の抽出とを関連付けて、行の抽出結果を段の抽出処理に反映させ、また段の抽出結果を行の抽出に反映させる手段を持つことである。

【0025】本発明を用いると、行方向の確定した行を基本としつつ、行抽出と段抽出の相互作用によりお互いの結果を高めあうことによって、従来技術ではなし得なかつた横行縦行が混在し、さらに段中の文字行が細分化されている不定形の段や行間よりも段間が狭い部分が存在するような複雑に入り組んだ文書構造を持つ文書画像に対しても、行および段の抽出を高精度に行うことができる。

【0026】以上の各処理手段を計算機によって実現するためのプログラムは、計算機が読み取り可能な可搬媒体メモリ、半導体メモリ、ハードディスクなどの適当な記憶媒体に格納することができる。

【0027】

【発明の実施の形態】以下、図面を参照して本発明の実施の形態を詳しく説明する。

【1】ハードウェア構成

図2は、本発明を適用するハードウェア構成例を示す。図1に示す文書画像認識装置100は、ハードウェアとしては、例えば図2に示すようなCPU20、メモリ21、キーボードその他の入力装置22、ディスプレイその他の出力装置23、ハードディスクその他の外部記憶装置24、画像を読み取るイメージスキャナ25およびこれらを接続するバス26等から構成される。図1に示す各処理手段の機能は、図2に示すメモリ21に格納されたプログラムをCPU20が実行することにより実現される。

【0028】【2】文書画像傾き検出部

図3は、本発明に係る文書画像傾き検出部1の基本構成例を示す。本発明による文書画像傾き検出部1では、文書画像2を入力し、傾きを検出して傾き角度3を出力する。文書画像傾き検出部1は、文字成分抽出手段11、行候補抽出手段12、行信頼度推定手段13、行抽出手段14、傾き推定手段15を持つ。

【0029】図4は文書画像傾き検出部1の処理フローチャート、図5は文字成分の抽出を説明する図、図6および図7は行候補の抽出を説明する図、図8は行信頼度の推定を説明する図である。

【0030】図4に示すフローチャートに従って、文書画像傾き検出部1の処理を詳しく説明する。まず、図4に示すステップS1で文書画像2を入力すると、文字成分抽出手段11は、文書画像2を参照し、文字を構成する成分の集合を抽出する。文字成分の抽出では、文書画像2中の黒画素の連結成分の集合をラベリングなどの手

法により抽出する（ステップS2）。このときに、各連結成分の外接矩形を求めておき、以下の処理では外接矩形の座標値のみを用いることにし、計算量の削減を図る。

【0031】次に、連結成分の外接矩形で互いに重なっているものを統合する。この処理を重なり矩形統合と呼ぶ。重なり矩形統合の結果、重なり矩形の集合が得られる。次に、重なり矩形の大きさのヒストグラムを取るなどして、図、表などの領域に対応した極端に大きな重なり矩形を除去することにより、文字成分に対応した文字成分外接矩形の集合を出力する。

【0032】図5（A）に示すように、例えば「日」という文字は、黒画素の一つの連結成分40が一つの文字成分として抽出される。図5（B）に示すように、

「日」という文字は、二つの黒画素の連結成分41、42から構成されるため、二つの文字成分として抽出される。図5（C）に示すように「お」という文字は、二つの黒画素の連結成分43a、43bから構成されるが、これらの外接矩形44a、44bをそれぞれ求めると、これらは互いに重なっている矩形であることが分かる。そこで、重なり矩形統合を行うことにより、二つの連結成分43a、43bが一つの文字成分として抽出される。

【0033】行候補抽出手段12は、図6に示す文書イメージ中の文字成分の配置50の例のような文字成分の集合を参照し、横および縦の行候補の集合を抽出する（ステップS3、S4）。なお、文字成分51のように、一つの文字成分が横行および縦行の二つの行候補に属することもある。

【0034】行候補抽出手段12は、具体的には、文字成分外接矩形同士の隣接関係から横行および縦行の候補集合を出力する。横行の場合について説明すると、横行に含まれている文字成分の一部は、下辺が行の基準線にほぼ沿っていることを考慮し、文字成分の外接矩形の下辺が横方向の一定の範囲内に存在するかを順次探索することにより、横行候補を抽出する。

【0035】今、図7（A）に示す文字成分の外接矩形60に隣接する横方向の文字成分を探索する場合を考える。文字成分の外接矩形60の下辺62に着目して、この下辺62の右端から右方向へ所定の角度（例えば±5度）の範囲内で、かつ、ある距離k・Xwの範囲内（ただし、kは所定値、Xwは外接矩形のX方向の幅）に、隣接する文字成分の外接矩形61の下辺63が存在するかどうかを調べる。もし、この範囲内に文字成分の外接矩形61の下辺63が存在すれば、外接矩形60から外接矩形61へ右リンクを張る。この探索を各外接矩形に対して繰り返す。なお、各外接矩形からの右リンクは、0本か1本のいずれかとし、右リンクが一つの外接矩形から枝分かれしないようにする。この結果、例えば図7（B）に示すように、横方向への右リンクが張られた外

接矩形群が求まるので、右リンクを辿ることにより、横行候補の集合を抽出する。

【0036】縦行候補についても同様に、それぞれの文字成分外接矩形に対して右辺が下隣にある文字成分に下リンクを張り、下リンクを辿ることにより求める。行信頼度推定手段13は、各々の行候補に対して信頼度を推定する（ステップS5）。まず、各々の行候補に対して行らしさを定量的に計算する。例えば、その行候補に含まれている文字成分の数が一定数以上あるという行の長さによる量、文字成分同士が比較的密に配置されているという近接性による量、文字成分の大きさがほぼ同じくらいであるなどの同質性による量などから行らしさを定量化できる。

【0037】次に、それぞれの文字成分外接矩形に対して、属する横行候補および縦行候補が一つずつ存在する可能性があるが、横行の行らしさと縦行の行らしさを比較してより行らしい方向の行に属する可能性が高いとして、その文字成分の方向フラグを横または縦にセットする。この結果から、それぞれの行候補に対して、その行候補に属する文字成分のうち、その行候補の行方向と一致する方向フラグを持つものの比率を求め、その行候補の信頼度とする。

【0038】例えば、図8（A）～（D）に示すように、各文字成分には、行候補抽出手段12の処理において、隣接する文字成分が見つからなかった文字成分70、縦行候補にのみ含まれる文字成分71、横行候補にのみ含まれる文字成分72、横行の縦行のいずれにも含まれる文字成分73等が存在する。これらの各文字成分に対して、前述した行らしさの尺度を利用して、行らしさを定量的に計算し、「横」または「縦」の方向フラグをセットする。なお、単なるフラグではなく、不明のものについて、縦0.5、横0.5というように確率的な数値を付与してもよい。この結果から、方向一致の比率を計算し、それを行信頼度とする。例えば、横行としての行信頼度は、図8（E）に示すように、着目している横行候補に対して、「横」の方向フラグがセットされている文字成分の比率を計算することにより求められる。

【0039】行抽出手段14は、行信頼度に基づいて確度の高い行の集合を抽出する（ステップS6）。ここでは、一定の信頼度以上の行信頼度を持つ行候補を確度の高い行として出力する。

【0040】傾き推定手段15は、確度の高い行の集合の文字成分の配置を用いて傾きを推定する（ステップS7）。確度の高い行の集合に属する文字成分外接矩形の基準辺（横行の場合は下辺、縦行の場合は右辺）の座標値から、複数直線の当てはめによる最小自乗法を用いて傾きの推定を行う。

【0041】以上のような文書画像傾き検出部1を用いると、横行と縦行とが混在する文書に対しても高速かつ高精度に傾きを検出することができる。検出した傾き角

度3によって、文書画像2の傾き補正を行い、次のレイアウト解析処理に進む。

#### 【0042】[3] レイアウト解析部

図9は、本発明に係るレイアウト解析部5の基本構成例を示す。本発明による文書画像認識装置100におけるレイアウト解析部5では、補正後の文書画像の基礎要素の集合4を入力とし、行抽出および段抽出をしてレイアウト解析結果6を出力する。このため、レイアウト解析部5は、基本行抽出手段8と行・段相互抽出手段9とを持つ。

【0043】さらに、基本行抽出手段8は、基礎要素集合の関連付けをもとに、横方向および縦方向の行方向を持つ行要素を生成する行要素生成手段81と、各々の行要素に対して信頼度を推定し、信頼度に基づいて行要素を淘汰する行要素淘汰手段82と、整合性のとれた行および文字集合を得るために行要素整合手段83とを持つ。また、行・段相互抽出手段9は、互いに関連付けられた段同士の統合を行う段要素統合手段91と、整合性のとれた段および行集合を得るために段要素整合手段92と、整合性のとれた行および文字集合を得るために行要素整合手段93とを持つ。

#### 【0044】(1) 基本行抽出手段

図10は、基本行抽出手段の処理概要を示す。基本行抽出手段8の入力は、例えば文書画像中の黒画素連結成分、あるいは文書画像中の黒画素の連結成分の外接矩形の重なり矩形というような、文書を構成する基礎要素の集合4であり、出力は、行要素、文字要素、基礎要素が互いに包含関係により関連付けられた木構造の集合である。

【0045】まず、ステップS10で、基礎要素の集合4を入力すると（ステップS10），行要素生成手段81は、各基礎要素を唯一含む文字要素を生成し（ステップS11），さらに文字要素の関連付けによる行要素を生成する（ステップS12）。具体的には、行要素生成手段81は、基礎要素の集合同士の関連付けをもとに、横方向および縦方向の行方向を持つ行要素を生成する。この基礎要素集合同士の関連付けは、近接性あるいは同質性に基づいて行う。

【0046】続いて、行要素淘汰手段82は、各々の行要素に対して信頼度を推定し、この信頼度に基づいて行要素を淘汰する（ステップS13）。さらに、行要素整合手段83は、整合性のとれた行集合と文字集合を得るために行要素集合の整合処理を行い（ステップS14），行要素、文字要素、基礎要素（重なり矩形要素）が互いに包含関係により関連付けられた木構造の集合を出力する。具体的には、以下のようないくつかの条件を満たす行要素集合および行要素の内部構造を得る。

【0047】1) すべての文字成分は唯一の行要素に属する。

50 2) 各々の行要素の行領域は互いに重ならない。ここ

で、行領域は各々の行要素に属する文字成分の外接矩形で定義される領域である。

【0048】3) 行方向の確定した行要素に属する文字要素は、行方向順に順序付けされ、行と垂直方向に重なる文字要素が存在しないように互いに統合されている。以下、具体例に従って基本行抽出手段8による行抽出についてさらに詳しく説明する。

【0049】図11に示すような文書画像の行・段抽出を行うものとする。図11に示す文書画像について、ラベリングによる黒画素連結成分の抽出処理および重なり矩形統合処理により、重なり矩形集合を基礎要素の集合として抽出したものを入力する(図10のステップS10)。入力した重なり矩形集合のそれぞれ一つを子として持つ文字要素の集合を生成し、これを基本文字要素の集合とする(ステップS11)。

【0050】図12は、このようにして生成された文書画像の基本文字要素の集合の例を示している。図12中に示す各矩形が文字要素を表している。次に、文字要素の関連付け(例えば近接性)による行要素の生成を行う(ステップS12)。ここでは、入力された基本文字要素の集合のそれぞれの文字要素に対して、横方向および縦方向に文字要素自身の大きさのしきい値倍以下の距離にあり、かつ、最も近接している文字要素を探索してリンクを張り、横方向および縦方向に近接した文字要素の列を、行要素として生成する。行要素には、横の行方向を持つもの(横行要素)と、縦の行方向を持つもの(縦行要素)の両方が存在する。多くの文字要素は、横行要素と縦行要素の両方の子となって、要素同士の干渉が多く発生する。

【0051】図13は、文字要素の近接性により生成した横行要素の外接矩形を表示した文書画像の例を示す。また、図14は、文字要素の近接性により生成した縦行要素の外接矩形を表示した文書画像の例を示す。

【0052】前処理として行われる文書画像傾き検出部1による傾き補正のために、傾きがほぼ0である文書画像を前提にできるため、近接条件に傾きによる広がりを持たせないことにする。また、近接条件の距離のしきい値は、異なる段に属するであろう行が統合されてしまわないために、傾き検出による傾き補正の場合より厳しく小さな値を採用する。

【0053】次に、行信頼度を用いた行要素の淘汰を行う(ステップS13)。文書画像傾き検出部1による処理と同様の手法を用いる。まず、生成した行要素の行信頼度を、属する文字要素の平均間隔、属する文字要素の数、属する文字要素のすべてを含む外接矩形による行領域の縦横比率をもとに算出する。次に、それぞれの文字要素に対して、関連する行要素の方向を文字要素の方向とする。関連する行要素が、横行要素と縦行要素の2種類存在する場合には、行信頼度の大きな行要素の方向を文字要素の方向とする。さらに、生成した行要素の強度

を、その行要素に含まれる文字要素のうち、行要素の行方向と同一の方向を持つ文字要素の比率とし、強度がしきい値未満の行要素を淘汰し、しきい値以上のものを信頼度の高い基本行要素として残す。

【0054】図15は、淘汰により生き残った信頼度の高い基本行要素を、外接矩形でもって表示した文書画像の例を示す。図15では、本来の横行および縦行の一部が、それぞれ正しい行方向を持つ行要素として生き残っており、かつ、異なる段に属する行の誤った統合が存在しないことがわかる。

【0055】次に、行要素整合手段83による行要素集合の整合処理(ステップS14)を行う。この行要素整合手段83の処理フローチャートを図16に示す。まず、行要素同士の領域干渉を解消する(ステップS31)。行要素同士の外接矩形は互いに接触している可能性がある。ここで、行要素の外接矩形とは、その行要素に含まれている文字要素の外接矩形を含む最小の矩形を意味する。行要素同士の外接矩形が接触している場合には、行要素間の干渉とみなされるので、それを解消する。

【0056】次に、行要素と領域干渉している文字要素の吸收仮説を生成する(ステップS32)。行要素の外接矩形と接触しているが、行要素の子として属していない文字要素が存在する可能性がある。そこで、それぞれの行要素に対してこのような文字要素を探索して、子として吸収する。ここでは、文字要素を子として吸収した行要素を、吸收仮説と呼んでいる。

【0057】続いて、文字要素の吸收仮説の要素干渉を検証することにより淘汰を行う(ステップS33)。上記のステップS32の処理の場合に、複数の行要素に接触している文字要素はノイズである可能性が高いので、行要素の子とせずに削除するのが望ましい場合もあるため、この検証を行う。

【0058】次に、行要素同士の領域干渉を解決する(ステップS34)。上記ステップS32およびS33の処理により、結果的にある行要素の新たな子となった文字要素が存在する場合には、行要素の外接矩形が膨張する可能性がある。したがって、この場合には、行要素に対し、ステップS31～S33の処理を再度繰り返す。

【0059】さらに、行要素に属する文字要素の整頓を行う(ステップS35)。それぞれの行要素の外接矩形に含まれる文字要素がすべて子となった状態を構成できたらば、次に行方向と垂直方向に重なっている文字要素は一つにまとめ、テキスト認識実行時の文字切り出し候補の可能性を限定するのが望ましい。このときに、まとめられた複数の文字要素の子となっている重なり矩形要素の基礎要素集合を一つに統合して、新たな統合された文字要素の子とする。

【0060】以上の行要素の整合処理を、図17に示す

画像例の1行のイメージに注目して説明する。図17の画像A中の矩形は、行要素淘汰後に得られた信頼度の高い基本縦行要素である。図17の画像B中の矩形列は、画像Aの縦行要素の子となっている文字要素である。行要素の外形矩形に含まれていない文字要素が二つ存在することが分かる。画像中の「や」と「し」の文字要素である。

【0061】その後、画像の行要素同士の領域干渉の解消・行要素と領域干渉している文字要素の吸収仮説を生成／淘汰し、図17の画像Cのように、行要素の外形矩形と接触している文字要素をすべて子として取り込む。さらに行要素同士の領域干渉の解決を経て、行要素に属する文字要素の整頓を行った結果が、図17の画像Dである。画像Dにより、行方向に垂直な方向（この場合には横方向）に射影した場合に重なっている文字要素が統合されて、まとめられたことが分かる。

【0062】次に、行要素と近接した文字要素の行要素への統合処理を行った結果が、図17の画像Eである。この場合、全体が一つの行として統合されている。ここで、図17の画像Fは、統合後の行要素の子となっている文字要素集合を示している。

【0063】以上の処理を行った結果を図18に示す。この段階で、すべての文字要素がある行要素の子となり、行要素、文字要素、基礎要素が互いに包含関係により関連付けられた木構造の集合が、中間的文書構造として構築できることになる。

【0064】以上の行抽出技術を用いた結果得られた行要素の集合には、図18の例からも明らかなように、段を構成した場合に行方向に近接している複数の行要素が含まれている。この意味で、この段階で得られた行要素は、段の構成要素としての完全な行には対応していない。完全な行を構成するためには、段抽出との連携を経て、同一の段に属する行要素をさらに統合するという、行抽出と段抽出との連携した処理を行うことが必要である。

#### 【0065】(2) 行・段相互抽出手段

行・段相互抽出手段9は、互いに関連付けられた段同士の統合と、整合性のとれた段および行集合を得るために段要素集合の整合と、整合性のとれた行および文字集合を得るために行要素集合の整合処理を実行する。さらに、行・段相互抽出手段9が、段同士の統合を繰り返し実行することにより、行抽出結果と段抽出結果が互いに影響を与えあい、徐々に高精度な行・段抽出処理が行われるようにする。

【0066】図19は、行・段相互抽出手段9の処理の流れの概要を示す。行・段相互抽出手段9の入力は、基本行抽出手段8により得られた行要素を頂点とする木構造の集合（行要素、文字要素、基礎要素が互いに包含関係により関連付けられた木構造の集合）であり、出力は、段要素を頂点とする木構造の集合（段要素、行要素、文字要素、基礎要素が互いに包含関係により関連付けられた木構造の集合）である。

【0067】行要素を頂点とする木構造の集合を入力すると（ステップS40），まず各行要素を唯一含む段要素の生成を行う（ステップS41）。この段要素について、以下のステップS42～S46を繰り返し実行することにより、段要素を統合していく。段要素間の関連付け（ステップS42），関連付けられた段要素の統合（ステップS43）は、基本的には、基本行抽出手段8において、行要素の生成のため文字要素と行要素の集合を対象に行っていた処理と同様であり、この処理を、行要素と段要素の集合を対象にして行う。

【0068】次に、段要素集合の整合処理を行う（ステップS44）。この段要素集合の整合処理は、以下の条件を満たす段要素集合および段要素の内部構造を得ることを目的とした処理である。

【0069】1) すべての行成分は唯一の段要素に属する。

2) 各々の段要素の段領域は互いに重ならない。ここで、段領域は各々の段要素に属する行成分の外接矩形で定義される領域である。

【0070】3) 段方向の確定した段要素に属する行要素は、段方向順に順序付けされ、段方向と垂直方向に重なる行要素が存在しないように互いに統合されている。この段要素集合の整合処理を行った後に、再度、基本行抽出手段8において説明した行要素集合の整合処理を行う（S45）。これにより、段抽出結果が行抽出結果にフィードバックされ、行および段の抽出の精度が高められる。以上のステップS42～S45の処理を、新たな関連付けが発生しなくなるまで繰り返す（ステップS46）。

【0071】以上の処理において、段要素間の関連付けを、行方向、行垂直方向とそれぞれ変化させて繰り返し実行することにより、2次元方向の関連付けを回避し、整合などの計算量を削減することもできる。また、段要素間の関連付けを、関連付けの対象となる段要素の内部構造により動的に決定することもでき、これにより文書構造に応じた高精度な行・段抽出を促進できる。例えば、大きな段が少数存在する整頓された段を持つ文書構造の場合には、パラメータを固定しても十分である。しかし、少数の行からなる小さな段が多数存在する不定形の段を持つ文書構造、あるいは空白がセパレータの役割を果たす表形式の文書構造の場合には、状況に応じた仮説生成を行う手法が有効である。状況に応じた段要素同士の関連付けとして、以下の3種類が考えられる。

【0072】1) 関連付け対象となる段の大きさ（含む行の数）により、近接性判定距離の標準サイズとの比較を変化させる。これにより、大きな安定した段同士の過統合を防止しつつ、細分化された小さな段同士あるいは大きな段と小さな段との適切な統合を促進できる。

【0073】2) 関連付け対象となる段に含まれている行のサイズに比例して、近接性判定距離を設定する。これにより、含む文字要素の大きさに応じて適切な統合を行えるため、行サイズの大きい段同士の未統合、行サイズの小さい段同士の過統合、行サイズの大きく異なる段同士の過統合を大幅に削減することができる。

【0074】3) 段の拡張処理の段階に応じて、近接性判定距離の標準サイズとの比率を変化させる。行方向の段の拡張処理は2回繰り返されるが、1回目は段が未だ細分化されており統合を促進する必要がある一方、2回目はある程度安定した段が生成できているため、むしろ過統合を抑制する必要がある。そこで、1回目は標準サイズとの比率を高く設定して統合しやすくし、2回目は標準サイズとの比率を小さく設定して統合しにくくした。

【0075】具体的には、近接性判定距離Dは、 $k \cdot S$ で与えられる。ここで、Sは、上記2)の処理で適応的に決定される標準サイズ、kは、上記1)および3)の処理で決定される係数である。上記1)および3)の処理を実現するために、例えば、段の統合を行垂直方向、行方向、行垂直方向、行方向と、計4回繰り返し、行方向の段統合において、図20に示すような係数kを用いると、高精度な段抽出が可能となり、段集合の整合処理の結果として高精度な行抽出も行うことができる。なお、図20では、「小さい段」は、それに含む行が1行のみの段、「中くらいの段」は、それに含む行が2行で、含む文字総数が100文字以内の段、「大きい段」は、それに含む行が2行で、含む文字総数が101文字以上、あるいは3行以上の段と、定義している。

【0076】また、段要素集合同士の関連付けは、近接性あるいは同質性に基づいて行うことができる。同質性を用いることにより、段間の方が行間よりも狭い場合でも精度よく段を分離することが可能となる。

【0077】さらに、段要素集合同士の関連付けに用いる要素関連付けパラメータを、処理回数に応じて変化させることにより、行・段の統合を進め、後半においては段の過統合を抑制するといった柔軟な処理を行い、行・段抽出の一層の高精度化を進めることができる。

【0078】文書画像例の段抽出結果を図21に示す。図22は、行・段相互抽出手段9により得られた段中の行要素を示している。これにより、同じ段に属する行は、行方向の射影に対して互いに重なりのない完全な行要素が抽出できていることが分かる。

【0079】

【実施例】以下に、図23ないし図25を用いて、本発明の実施例を説明する。横行縦行が混在する文書である新聞文書の文書画像を対象事例として、図23に示すような、新聞紙面から400d p iで画像取り込みを行った二値化文書画像を対象とする。なお、図23の文書画像は縮小表示している。また、左上部分には、写真が印

刷されているが、ここでは認識対象にはならないので空白で置き換えてある。

【0080】図24は、図23の文書画像についての行抽出処理の結果の例を示す。この文書画像についての行抽出処理の抽出精度は100%であった。図25は、最終的な段抽出処理の結果の例を示す。この文書画像についての段抽出処理の抽出精度も100%であった。

【0081】

【発明の効果】以上説明したように、本発明によれば、従来技術では困難であった横書き（横行）、縦書き（縦行）が混在した文書画像の傾きを高速かつ高精度で検出することができ、さらに段中の文字行が細分化されている不定形の段や、行間よりも段間が狭い部分が存在する複雑に入り組んだ文書構造を持つ文書画像中の文字行および段の抽出を高精度に行うことができるようになる。

【図面の簡単な説明】

【図1】本発明の基本構成例を示す図である。

【図2】本発明を適用するハードウェア構成例を示す図である。

【図3】本発明に係る文書画像傾き検出部の基本構成例を示す図である。

【図4】文書画像傾き検出部の処理フローチャートである。

【図5】文字成分の抽出を説明する図である。

【図6】行候補の抽出を説明する図である。

【図7】行候補の抽出を説明する図である。

【図8】行信頼度の推定を説明する図である。

【図9】本発明に係るレイアウト解析部の基本構成例を示す図である。

【図10】基本行抽出手段の処理概要を示す図である。

【図11】処理対象の文書画像の例を示す図である。

【図12】基本文字要素の集合の例を示す図である。

【図13】横行要素の外接矩形を表示した文書画像の例を示す図である。

【図14】縦行要素の外接矩形を表示した文書画像の例を示す図である。

【図15】淘汰により生き残った信頼度の高い行要素を外接矩形表示した文書画像の例を示す図である。

【図16】行要素整合手段の処理フローチャートである。

【図17】行要素の整合処理により処理した画像の例を説明する図である。

【図18】基本行抽出手段により抽出した結果の例を示す図である。

【図19】行・段相互抽出手段の処理の流れの概要を示す図である。

【図20】近接性判定距離の標準サイズとの比率を変化させる係数kの例を示す図である。

【図21】行・段相互抽出手段により抽出した結果の段要素の例を示す図である。

【図22】行・段相互抽出手段により抽出した結果の行要素の例を示す図である。

【図23】本発明の実施例の評価を説明するための図である。

【図24】本発明の実施例の評価を説明するための図である。

【図25】本発明の実施例の評価を説明するための図である。

【図26】従来技術の問題を説明するための図である。

【符号の説明】

100 文書画像認識装置

1 文書画像傾き検出部

11 文字成分抽出手段

12 行候補抽出手段

13 行信頼度推定手段

14 行抽出手段

15 傾き推定手段

2 文書画像

3 傾き角度

4 基礎要素の集合

5 レイアウト解析部

6 レイアウト解析結果

8 基本行抽出手段

81 行要素生成手段

10 82 行要素淘汰手段

83 行要素整合手段

9 行・段相互抽出手段

91 段要素統合手段

92 段要素整合手段

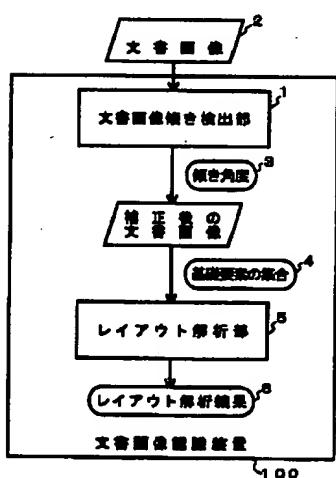
93 行要素整合手段

【図1】

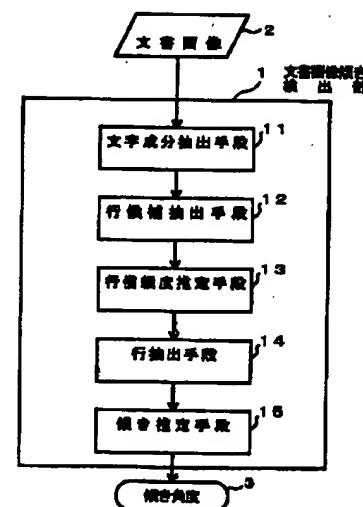
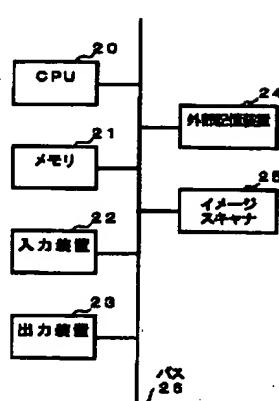
【図2】

【図3】

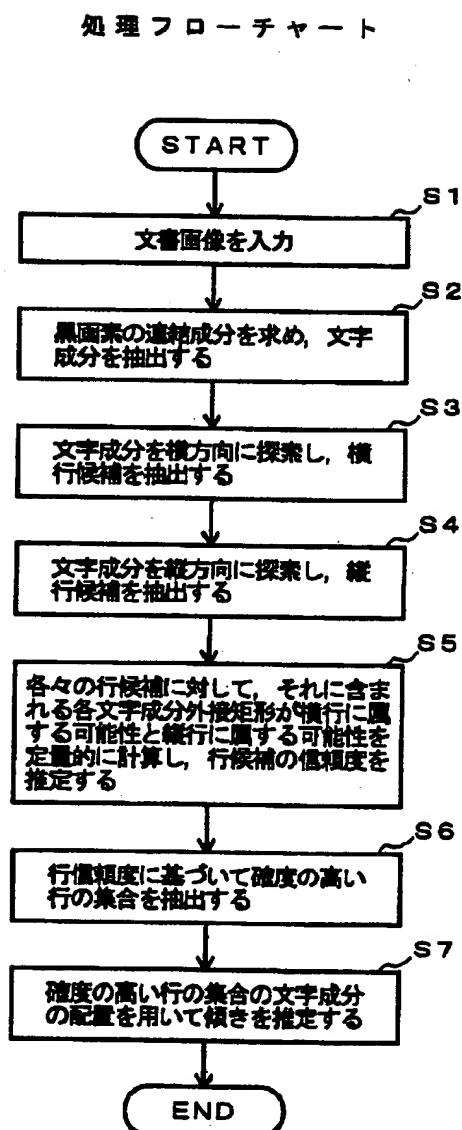
本発明の基本構成例



ハードウェア構成例

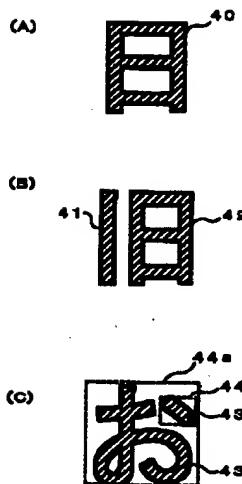


【図4】



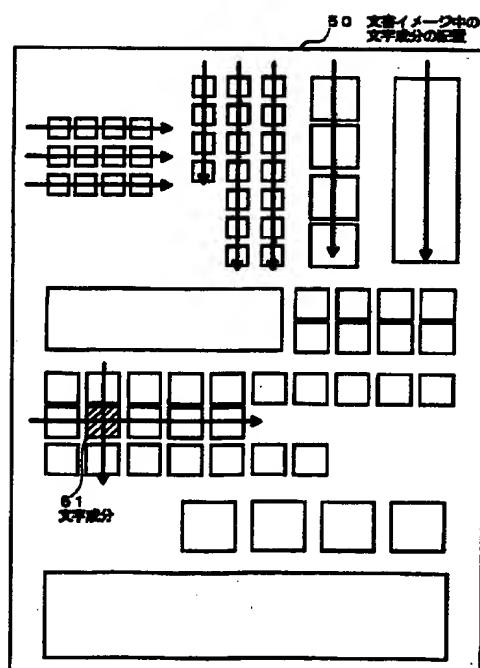
【図5】

文字成分の抽出

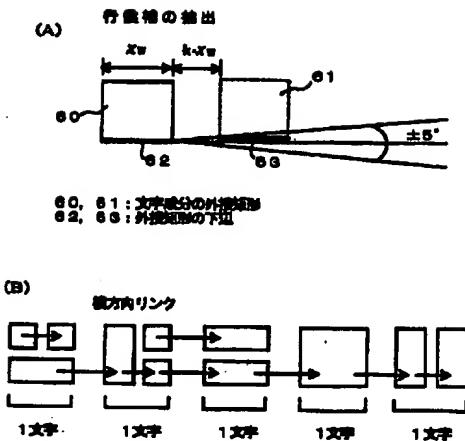


【図6】

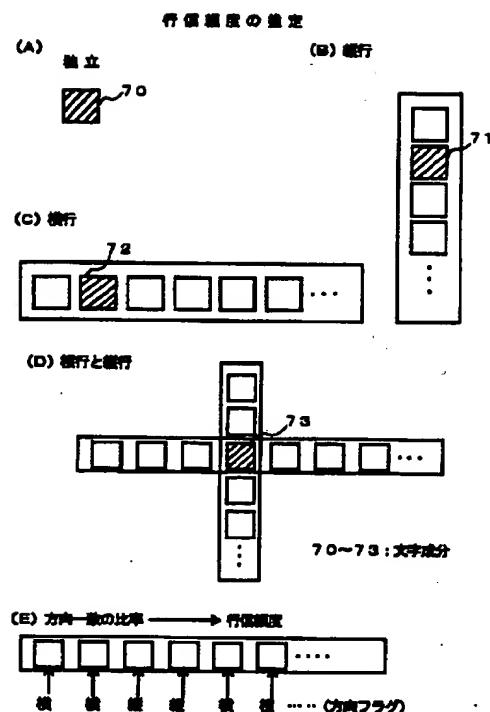
行候補の抽出



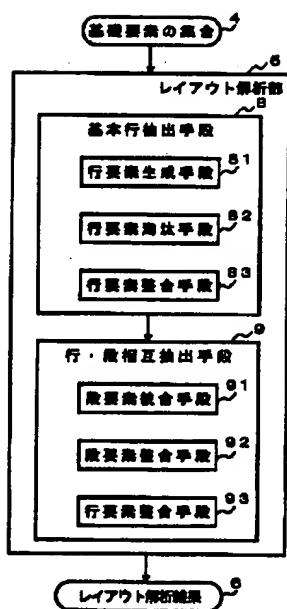
(図7)



(図8)



〔圖9〕



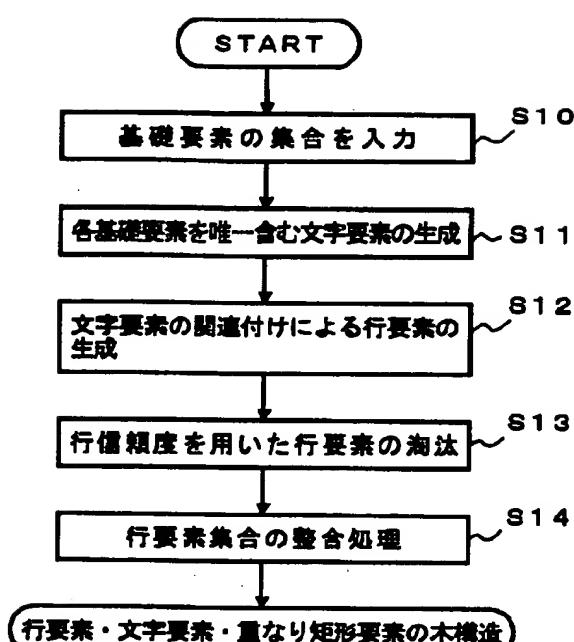
### 【図11】

10月からの酒税法の改正＝ホーワード率廃止を前に、ウイスキーと焼酎（しょうちゅう）の太字メニューの改定が立ててきた。酒類の種類が豊かなもウイスキーメニューは、1000円前後の低価格

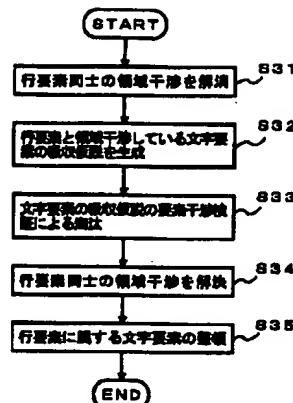
量を相次いで投入する。一方、純粋メーカーはアリで、アルコール度数が低くて価格が安い商品を出したり、ボトルを工夫したりして、競争を行っていく。

## 度数半分の新商品

【図10】

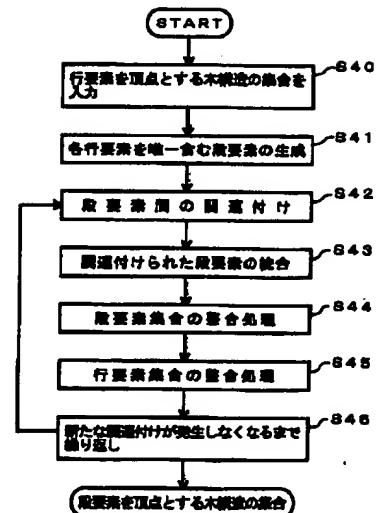
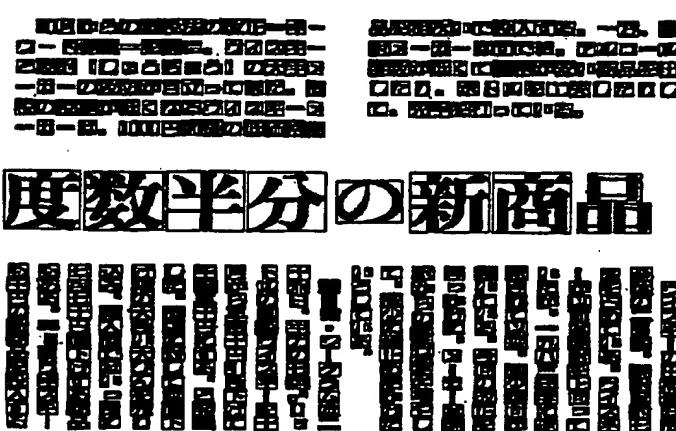


【図16】



【図19】

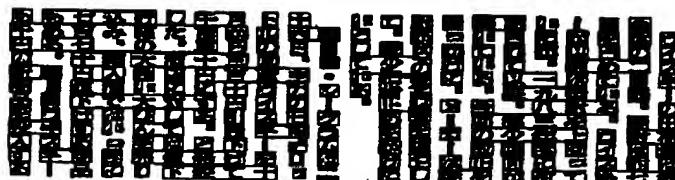
【図12】



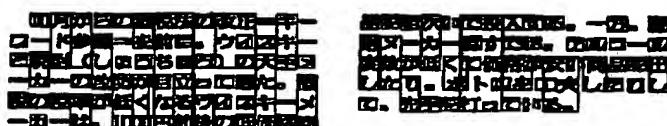
【図13】



## 度数半分の新商品



【図14】



## 度数半分の新商品



【図20】

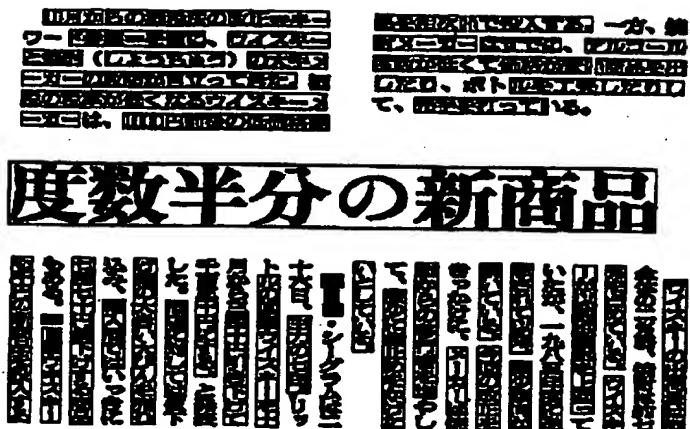
(A) 1回目の行方向の段の統合時

	小さい段	中くらいの段	大きい段
小さい段	3.0	2.0	1.0
中くらいの段	2.0	2.0	1.0
大きい段	1.0	1.0	1.0

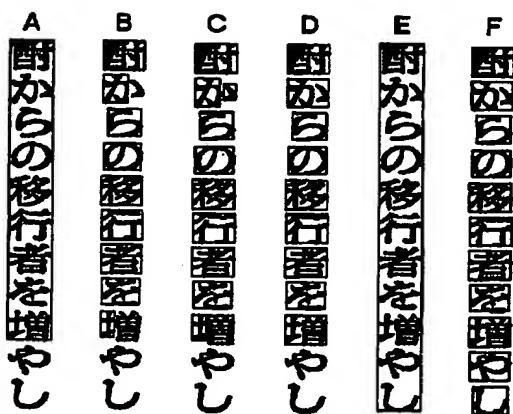
(B) 2回目の行方向の段の統合時

	小さい段	中くらいの段	大きい段
小さい段	1.0	1.0	0.5
中くらいの段	1.0	1.0	0.5
大きい段	0.5	0.5	0.5

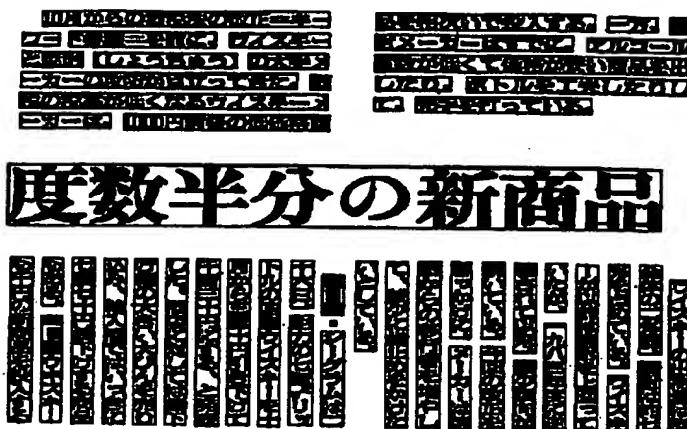
[図15]



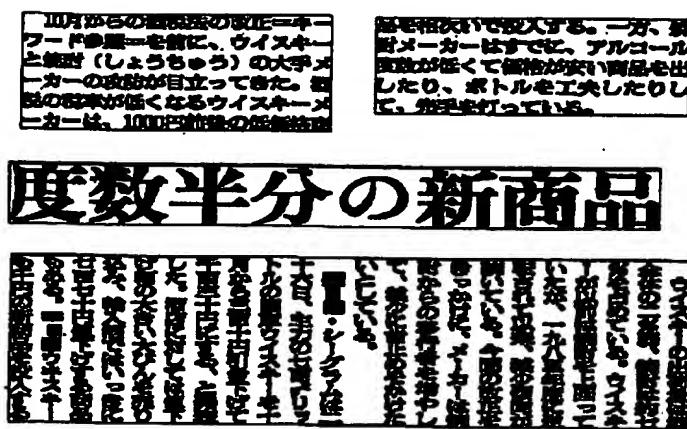
〔図17〕



[図18]



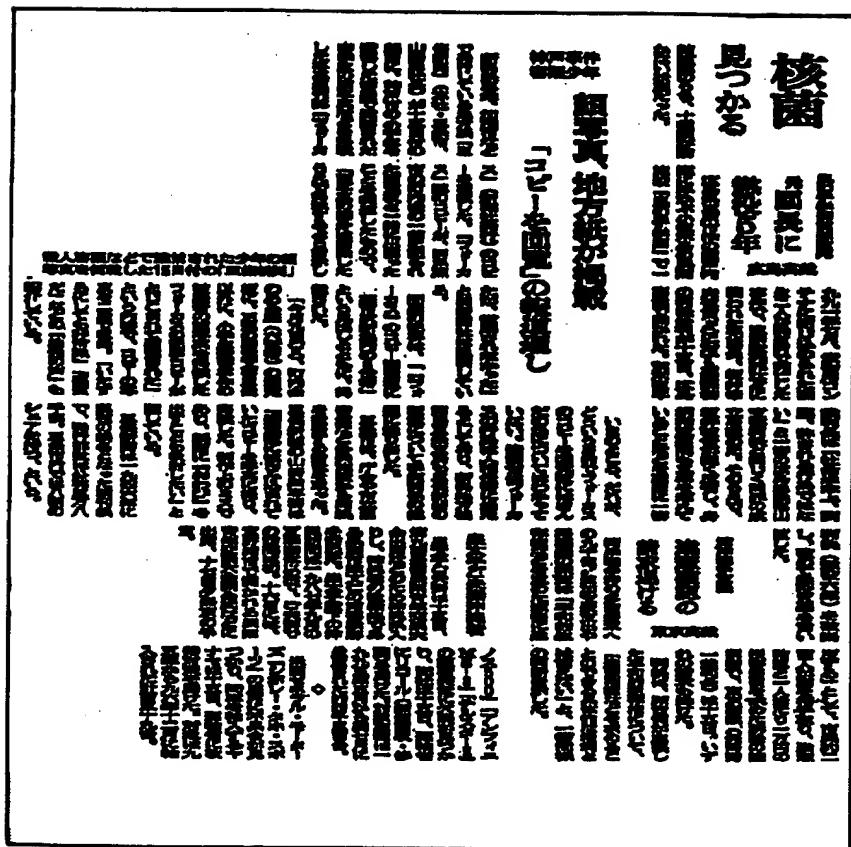
[図21]



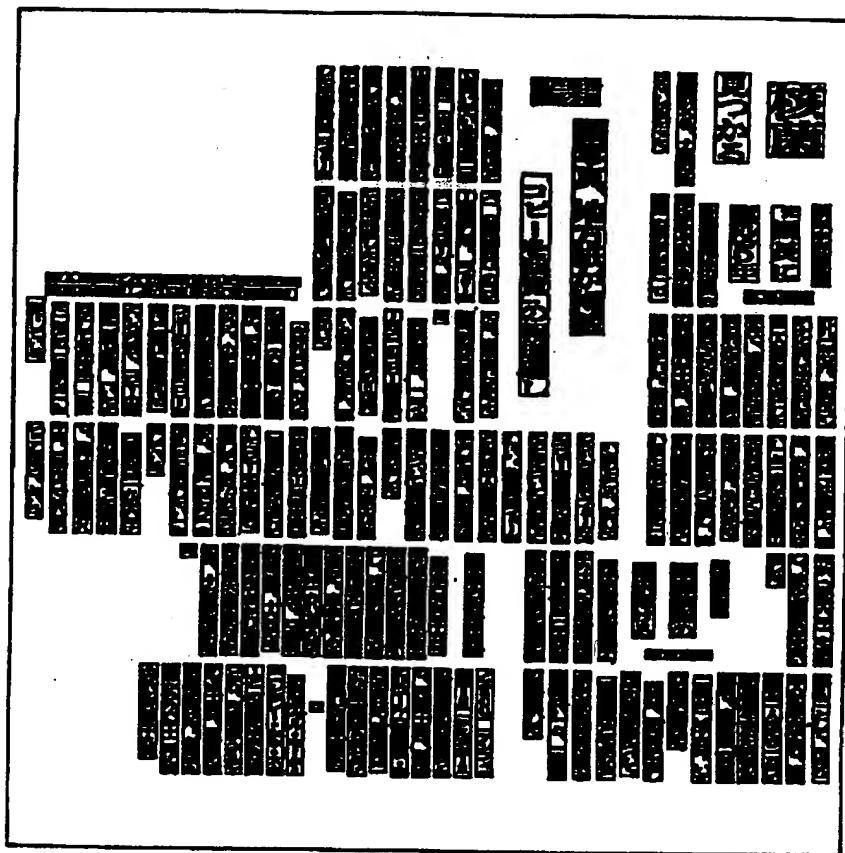
【図22】



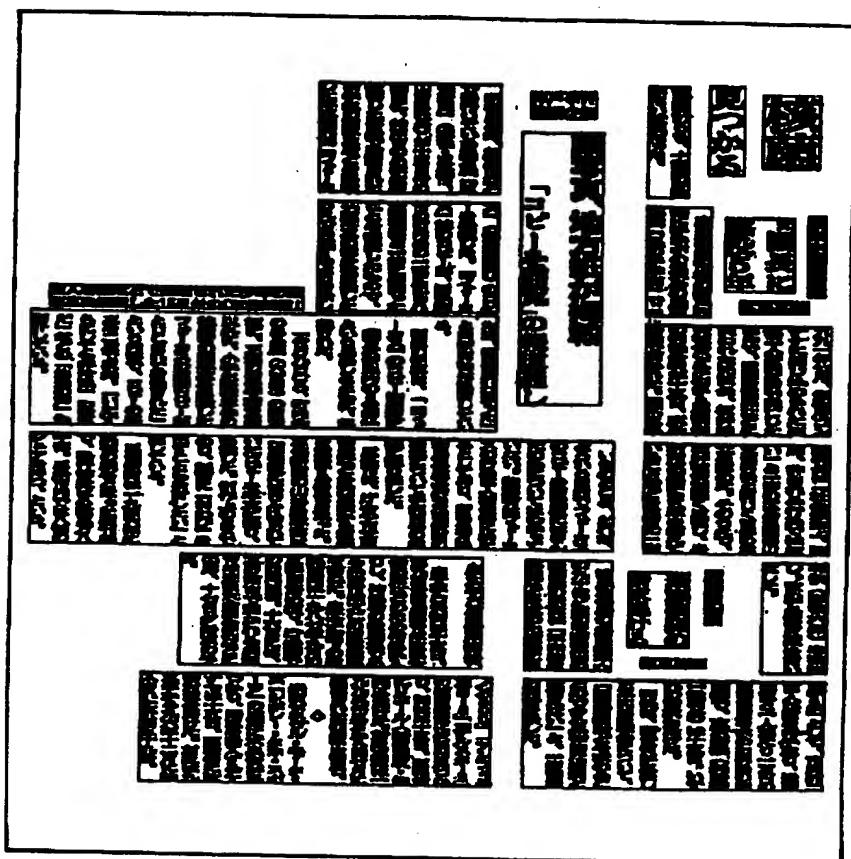
[図23]



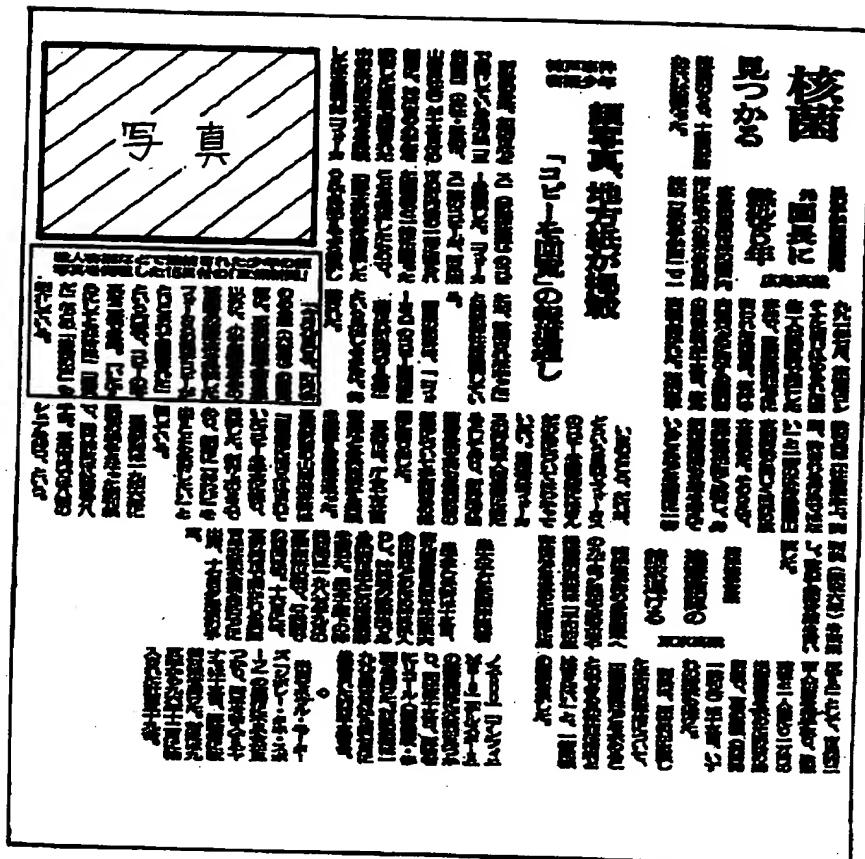
【図24】



【図25】



【図26】



**This Page is Inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record**

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- BLACK BORDERS**
- IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- FADED TEXT OR DRAWING**
- BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- SKEWED/SLANTED IMAGES**
- COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- GRAY SCALE DOCUMENTS**
- LINES OR MARKS ON ORIGINAL DOCUMENT**
- REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- OTHER:** \_\_\_\_\_

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.**